Randomized controlled trials in very preterm infants: does inclusion in the study result in any long-term benefit?

Rüegger, Christoph M ; Kraus, Andrea ; Koller, Brigitte ; Natalucci, Giancarlo ; Latal, Beatrice ; Waldesbühl, Etienne ; Fauchère, Jean-Claude ; Held, Leonhard ; Bucher, Hans Ulrich

Abstract: BACKGROUND Since the introduction of randomized controlled trials (RCT) in clinical research, there has been discussion of whether enrolled patients have worse or better outcomes than comparable non-participants. OBJECTIVE To investigate whether very preterm infants randomized to a placebo group in an RCT have equivalent neurodevelopmental outcomes to infants who were eligible but not randomized (eligible NR). METHODS In the course of an RCT investigating the neuroprotective effect of early high-dose erythropoietin on the neurodevelopment of very preterm infants, the outcome data of 72 infants randomized to placebo were retrospectively compared with those of 108 eligible NR infants. Our primary outcome measures were the mental (MDI) and psychomotor (PDI) developmental indices of the Bayley Scales of Infant Development II at 24 months of corrected age. The outcomes of the two groups were considered equivalent if the confidence intervals (CIs) of their mean differences fitted within our ±5-point margin of equivalence. RESULTS Except for a higher socioeconomic status of the trial participants, both groups were balanced for most perinatal variables. The mean difference (90% CI) between the eligible NR and the placebo group was -2.1 (-6.1 and 1.9) points for the MDI and -0.8 (-4.2 and 2.5) points for the PDI. After adjusting for the socioeconomic status, maternal age and child age at follow-up, the mean difference for the MDI was -0.5 (-4.3 and 3.4) points. CONCLUSIONS Our results indicate that the participation of very preterm infants in an RCT is associated with equivalent long-term outcomes compared to non-participating infants. © 2014 S. Karger AG, Basel.

DOI: https://doi.org/10.1159/000362784

Original Paper

Neonatology

Neonatology 2014;106:114–119
DOI: 10.1159/000362784

Received: January 7, 2014
Accepted after revision: April 9, 2014
Published online: June 20, 2014

Randomized Controlled Trials in Very Preterm Infants: Does Inclusion in the Study Result in Any Long-Term Benefit?

Christoph M. Rüegger[a] Andrea Kraus[b] Brigitte Koller[a] Giancarlo Natalucci[a, c] Beatrice Latal[c] Etienne Waldesbühl[a] Jean-Claude Fauchère[a] Leonhard Held[b] Hans Ulrich Bucher[a]

[a]Division of Neonatology, University Hospital Zurich, [b]Division of Biostatistics, University of Zurich, and [c]Child Development Center, University Children's Hospital, Zurich, Switzerland

Key Words
Preterm infants · Trial effect · Long-term outcome · Randomized controlled trial

Abstract
Background: Since the introduction of randomized controlled trials (RCT) in clinical research, there has been discussion of whether enrolled patients have worse or better outcomes than comparable non-participants. ***Objective:*** To investigate whether very preterm infants randomized to a placebo group in an RCT have equivalent neurodevelopmental outcomes to infants who were eligible but not randomized (eligible NR). ***Methods:*** In the course of an RCT investigating the neuroprotective effect of early high-dose erythropoietin on the neurodevelopment of very preterm infants, the outcome data of 72 infants randomized to placebo were retrospectively compared with those of 108 eligible NR infants. Our primary outcome measures were the mental (MDI) and psychomotor (PDI) developmental indices of the Bayley Scales of Infant Development II at 24 months of corrected age. The outcomes of the two groups were considered equivalent if the confidence intervals (CIs) of their mean differences fitted within our ±5-point margin of equivalence. ***Results:*** Except for a higher socioeconomic status of the trial participants, both groups were balanced for most perinatal variables. The mean difference (90% CI) between the eligible NR and the placebo group was –2.1 (–6.1 and 1.9) points for the MDI and –0.8 (–4.2 and 2.5) points for the PDI. After adjusting for the socioeconomic status, maternal age and child age at follow-up, the mean difference for the MDI was –0.5 (–4.3 and 3.4) points. ***Conclusions:*** Our results indicate that the participation of very preterm infants in an RCT is associated with equivalent long-term outcomes compared to non-participating infants.

Introduction

Sixty-five years ago, the implementation of randomized controlled trials (RCT) ushered in a new era of medical research [1]. This type of scientific experiment has since been accepted as the gold standard for clinical trials and the most reliable method for evidence-based decision-making [2, 3]. For approximately the same period,

KARGER

E-Mail karger@karger.com
www.karger.com/neo

1661–7800/14/1062–0114$39.50/0

Christoph M. Rüegger, MD
Division of Neonatology, University Hospital Zurich
Frauenklinikstrasse 10
CH–8091 Zurich (Switzerland)
E-Mail christoph.rueegger@usz.ch

there has been widespread discussion about whether enrolled patients have worse or better outcomes than comparable non-participants. On the one hand, critical voices emphasize the potential risks to which patients are exposed when participating in a clinical trial. On the other hand, an increasing number of researchers and patients believe that participating in clinical research is neither risky nor dangerous but beneficial and desirable instead. However, there is limited evidence that such a positive 'trial effect' exists.

Previous research comparing the patient outcomes within and outside an RCT has largely focused on adults [4–6]. Although these studies varied in their conclusions, they all found that participating in a clinical trial more likely results in positive than negative effects. A major disadvantage of most of these systematic reviews is that they compared all patients who were treated in the trials with all patients who were not enrolled, regardless of the clinical intervention performed or differences in the patient populations. Consequently, the results are potentially influenced by a selection bias and at least two (i.e. treatment and protocol effect) of the five components that potentially favor clinical trials (table 1) [6].

In light of the sparse data available about a trial effect in preterm infants and no available data about its long-term effects, the objective of this study was to investigate whether very preterm infants enrolled in a clinical trial and randomized to placebo treatment have neurodevelopmental outcomes that are equivalent to those of infants who were eligible but not randomized (eligible NR).

Table 1. Components of the trial effect

Trial effect	Differences
Treatment effect	Treatment offered in a study is better than the current standard of care
Protocol effect	Strict adherence to well-defined protocols
Care effect	More extensive follow-up or nursing care
Hawthorne effect	Changes in patient or clinician behavior as a result of being under close observation in a trial
Placebo effect	Psychologically mediated benefits associated with the administration of a sham or simulated intervention

Methods

Our study group consisted of very preterm infants between $26^{0/7}$ and $31^{6/7}$ gestational weeks who were enrolled in an RCT lasting from September 2005 to January 2011 (www.clinicaltrials.gov; NCT00413946) [7]. The primary objective of this clinical trial was to determine whether the neurodevelopmental outcomes are improved in very preterm infants when they are administered early high-dose erythropoietin (EPO). Infants were excluded for any of the following reasons: born outside of a perinatal center, genetically defined syndromes, severe congenital malformations adversely affecting life expectancy or neurodevelopment, and infants who were a priori admitted for palliative care. Within the first 3 h of life, the infants were randomized to receive EPO or placebo (NaCl 0.9%) intravenously at 3, 12–18, and 36–42 h after birth. Other than the study treatment, all procedures and examinations were standard practice of care. At 24 months of age (corrected for prematurity), the mental and psychomotor development was assessed using the Bayley Scales of Infant Development II [8]. Of all infants enrolled in the EPO trial, only the infants born at our perinatal center and randomized to placebo were eligible for the present analysis. The infants with incomplete follow-up data at 2 years of age were excluded.

Our control group was restricted to concurrent preterm infants born at our perinatal center who would have met the eligibility criteria for the EPO trial. They were not enrolled in the EPO trial for any of the following reasons: failure to obtain parental consent, parents were not approached, language barriers, and failure to obtain written informed consent because of an emergency. Of the remaining infants, those with incomplete follow-up data at 2 years of age were excluded. Data on these eligible NR controls were collected prospectively and entered into a large database maintained by the Swiss Neonatal Network & Follow-Up Group. Socioeconomic status (SES) was estimated based on paternal occupation and maternal education [9]. The scores ranged from 2 (highest SES) to 12 (lowest SES). Our primary outcome measures were the mental (MDI) and psychomotor (PDI) developmental indices of the Bayley Scales of Infant Development II at 24 months of corrected age.

Ethics

The EPO trial was approved by the Ethical Committee of the Canton Zurich and by Swissmedic in Bern. Written informed consent was obtained from the parents of infants in the placebo group, ideally before birth. The parents of the eligible NR infants did not object that data from their infants would be used for scientific analysis after anonymization.

Statistics

Comparisons between groups were based on the Mann-Whitney test for continuous variables and the Fisher's exact test for nominal variables.

An equivalence test regarding the two primary outcomes was performed [10, 11]. The aim of such a test is to show that the two groups of interest do not differ by more than a prespecified margin of clinical relevance. This is different from classical superiority testing, where the aim is to show that the two groups differ. To perform an equivalence test at the level of 5%, a 90% confidence interval (CI) is constructed for the mean difference between the outcomes in the two groups. Equivalence is shown if the whole interval is between $-\Delta$ and Δ. As the width of the CI plays a crucial role here, we assessed the sample size with reference to precision rather than power.

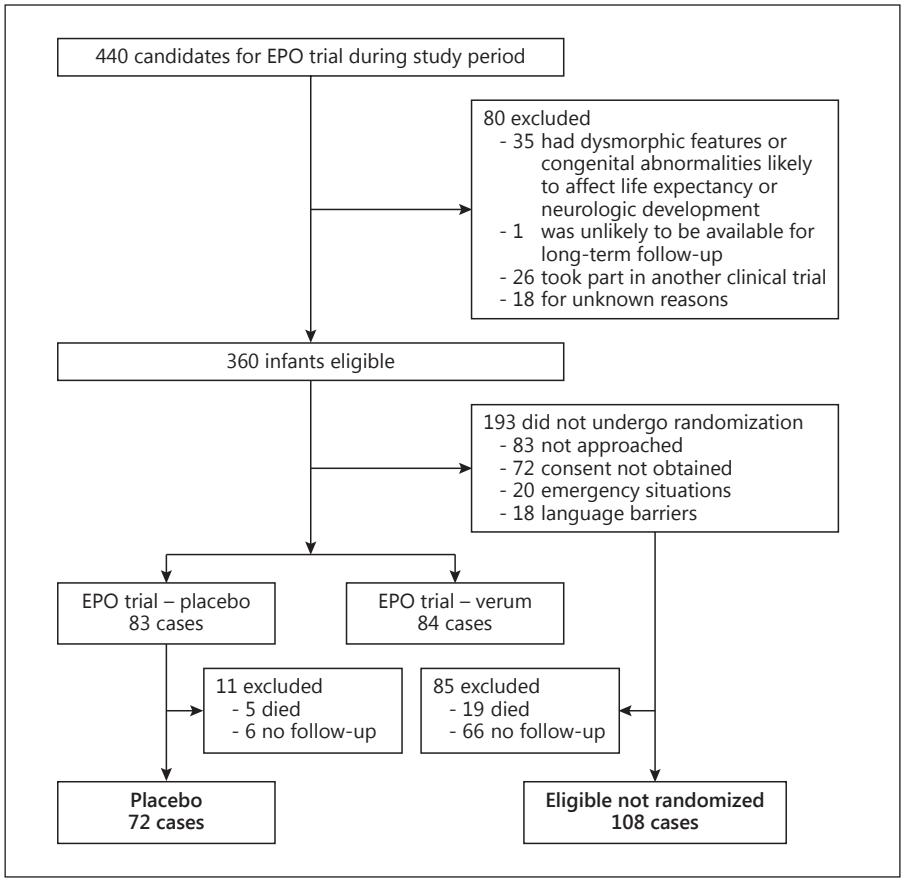

Fig. 1. Number of infants who were eligible for the study.

We set the margin of clinical relevance at $\Delta = 5$ points (0.3 standard deviations of the Bayley Scales of Infant Development II). We constructed unadjusted CIs based on the t test, as well as CIs adjusted for potential confounders based on linear regression. The sample size assessment was based on the assumption that the standard deviation of both MDI and PDI is 15 points. From the EPO trial, there were 83 infants in the placebo group and 193 eligible NR infants available (fig. 1). These sample sizes achieve a 90% CI for the unadjusted group difference of width 6.5 points (i.e. approximately two thirds of the width of the range between –5 and 5). This was considered to be sufficiently accurate, even after the possible exclusion of some infants due to loss to follow-up or death.

Results

A total of 180 infants born at our perinatal center were included in this study: 72 infants from the placebo group of the EPO trial, and 108 eligible but NR infants (fig. 1). The perinatal variables were evenly distributed, but SES and maternal age was higher in the placebo group compared to the eligible NR group (table 2).

The short-term outcome until discharge was similar in the two groups (table 3), but the eligible NR infants suffered more often from a low-grade intraventricular hemorrhage (IVH – grade I and II) than those infants randomized to placebo.

The follow-up rate at 2 years of age differed significantly between the two groups (placebo 92.3%, eligible NR 62.1%, $p < 0.001$). The demographics and short-term outcomes of the eligible NR infants with (n = 108) and without (n = 66) follow-up examinations are given in table 4, showing an increased rate of low-grade intraventricular hemorrhages and a trend towards a higher SES in the group of infants who were examined.

At the follow-up examination, the mean age ± standard deviation was 22.9 ± 1.7 months in the placebo group and 23.7 ± 2.4 months in the eligible NR group ($p < 0.001$). Neither the placebo nor eligible NR groups showed a significantly increased incidence of visual ($p = 0.49$) or hearing ($p = 0.40$) problems, and the rate of infants receiving physical, occupational or child psychiatry therapy ($p = 0.51$) was similar.

Table 2. Baseline characteristics

Characteristics	Placebo (n = 72)	Eligible NR (n = 108)	p value
Pregnancy complications[a]	34/72 (47.2)	46/108 (42.6)	0.54
Prenatal steroids	66/72 (91.7)	97/104 (93.3)	0.77
Male gender	43/72 (59.7)	53/108 (49.1)	0.17
Gestational age, weeks	29.5±1.5	29.5±1.7	0.84
Birth weight, g	1,236±354	1,222±335	1.00
Head circumference at birth, cm	27.3±2.3	27.0±2.2	0.61
Umbilical artery pH	7.33±0.1	7.31±0.1	0.36
Socioeconomic status score	5.2±2.3	6.0±2.7	0.04
Maternal age, years	33.1±6.0	31.1±5.7	0.03

Values are n (%) or means ± SD. [a] Chorioamnionitis, gestational diabetes, preeclampsia or HELLP syndrome.

Table 3. Short-term outcome until discharge

Outcome measures	Placebo (n = 72)	Eligible NR (n = 108)	p value
Mechanical ventilation, days	1.4±2.6	0.9±1.7	0.12
CPAP treatment, days	11.2±14.8	10.3±14.1	0.68
Supplemental oxygen, days	21.3±24.3	15.5±21.5	0.12
LOS of the survivors, days	55.0±19.8	56.0±25.7	0.69
Surfactant	24/72 (33.3)	33/108 (30.6)	0.74
IVH I and II	5/72 (6.9)	28/108 (25.9)	0.001
IVH III and IV	4/72 (5.6)	6/108 (5.6)	1.00
Sepsis	11/72 (15.3)	16/108 (14.8)	1.00
ROP 1–4	4/70 (5.7)	8/94 (8.5)	0.56
Necrotizing enterocolitis	4/72 (5.6)	4/107 (3.7)	0.72
BPD 1–3	9/72 (12.5)	10/106 (9.4)	0.62

Values are means ± SD or n (%). LOS = Length of stay; ROP = retinopathy of prematurity; BPD = bronchopulmonary dysplasia.

Table 4. Demographics and short-term outcomes of the eligible NR infants with and without follow-up examinations

Outcome measures	Eligible NR		p value
	with follow-up (n = 108)	without follow-up (n = 66)	
Socioeconomic status score	6.0±2.7	6.7±2.3	0.06
Maternal age, years	31.1±5.7	31.5±6.2	0.78
Mechanical ventilation, days	0.9±1.7	1.5±4.8	0.85
CPAP treatment, days	10.3±14.1	8.5±12.4	0.54
Supplemental oxygen, days	15.5±21.5	16.3±20.2	0.45
LOS of the survivors, days	56.0±25.7	51.5±19.6	0.33
Surfactant	33/108 (30.6)	16/66 (24.2)	0.39
IVH I and II	28/108 (25.9)	7/66 (10.6)	0.02
IVH III and IV	6/108 (5.6)	2/66 (3.0)	0.71
Sepsis	16/108 (14.8)	7/66 (10.6)	0.50
ROP 1–4	8/94 (8.5)	5/52 (9.6)	1.00
Necrotizing enterocolitis	4/107 (3.7)	1/66 (1.5)	0.65
BPD 1–3	10/106 (9.4)	3/65 (4.6)	0.37

Values are means ± SD or n (%). LOS = Length of stay; ROP = retinopathy of prematurity;BPD = bronchopulmonary dysplasia.

Our primary outcomes did not differ significantly between the two groups, but the unadjusted 90% CI of the mean MDI difference was –6.1 and 1.9 points (mean difference –2.1 points) and exceeded our margins of equivalence in favor of the infants in the placebo group (fig. 2). After adjusting for SES, maternal age, and child age at follow-up, the 90% CI of the mean MDI difference was within our predefined ±5-point margin of equivalence. The unadjusted 90% CI of the mean PDI difference was –4.2 and 2.6 points (mean difference –0.8 points) and did not materially change after the adjustment for SES, maternal age, and child's age at follow-up.

We also analyzed whether other perinatal risk factors were related to our primary outcomes. At 2 years of age, none of them had a significant impact on either MDI or PDI.

Discussion

The long-debated theory of a trial effect led us to infer that preterm infants enrolled in an RCT and randomized to placebo would have the same long-term outcomes than eligible NR controls. Our findings confirm this theory, as the participation in the placebo arm of the EPO trial was associated with equivalent long-term outcomes to receiving the same treatment outside the trial.

To date, only one study investigating a potential trial effect on the short-term outcome of very preterm infants has been performed [12]. First, the authors found a reduction in the median duration of mechanical ventilation from 6.2 days in the eligible NR group to 4.8 days in the placebo group (p = 0.008). Second, there was also a trend toward less frequent and less severe IVH in the trial participants. They concluded that sick newborns might benefit from participation in a RCT and speculated that a protocol effect and possibly the scrutiny from study personnel may have benefited infants in the placebo arm of the trial.

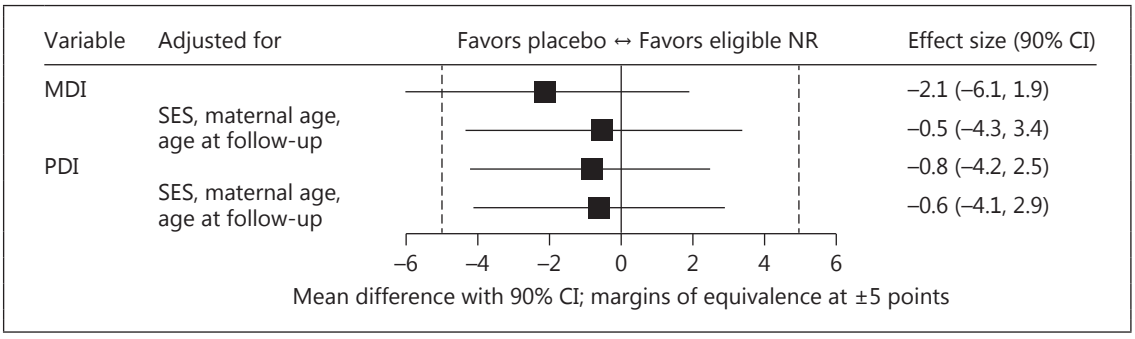

Fig. 2. Non-adjusted and adjusted long-term outcomes.

Our short-term outcomes were similar in both groups, except for the rate of low-grade IVHs. These bleedings can be difficult to identify [13] and might, therefore, be particularly sensitive to over- or underdiagnosing. An increased diagnostic quality based on an independent and blinded panel of experts supervising the correct grading of IVHs within an RCT might be the cause for the reported differences between the placebo and the eligible NR group.

Our main results aligned with those of the most recent published studies [14–16], although the concept of an 'inclusion benefit' differs in various ways between well-informed adults and very preterm infants. First, in addition to the study treatment, all procedures were the same in our placebo and eligible NR groups. An advantage arising from a better experimental treatment within the trial (treatment effect) can, therefore, be ruled out. Second, the treatment of very preterm infants in our neonatal intensive care unit is already largely standardized. Therefore, adhering to rigorous treatment guidelines, as outlined in a clinical trial (protocol effect), would not likely add value in such an already strictly controlled setting. Third, most neonatal intensive care units follow the policies of 'minimal handling' attempting to limit the number of interventions and their routine periodicity, which goes against the idea that trial participants might benefit from extra nursing cover or any additional examinations (care effect). Furthermore, our trial protocol did not require more comprehensive care for the randomized infants. Finally, there is little chance of intrinsic trial effects in very preterm infants.

For all these reasons, our main findings argue against a long-lasting trial effect in the most immature infants. However, we appreciate that parents, clinicians, and nurses may likewise develop placebo or Hawthorne effects that may result in altered decision-making. Further bias might arise from differences between parents who authorize the participation of their infant in an RCT and non-consenting parents. Our data confirm these subtle but important differences, as parents with a higher SES were more likely to allow participation of their infant in our trial. This finding is important because SES has also been the strongest confounder for the mental outcome, increasing the MDI by 2.0 points on average (95% CI 2.9 and 1.1) for every point towards a higher SES.

The variables SES, mother's age and child's age at follow-up were significantly imbalanced between the placebo and the eligible NR groups and at the same time, clinically relevant in terms of our primary outcomes. Consequently, all three variables were included in our linear regression models to eliminate bias from the analysis. Nevertheless, it remains unclear why the infants in the placebo group were significantly younger at follow-up.

Although such examinations at 2 years of corrected age are standard practice of care for all very low birth weight infants in Switzerland, there was a higher 'lost to follow-up' rate in the eligible NR group than in the placebo group. This difference alone is a significant finding, as better follow-up leads to earlier detection of developmental problems and, consequently, earlier initiation of therapeutic strategies. However, structured follow-up programs for premature infants are not yet implemented in all countries and exclusion from such preventive measures will more likely result in negative effects for infants who do not participate in the trial. Given that the infants within and outside the trial receive the same or similar treatment, it can likewise be stated that non-participation was not correlated with an adverse outcome. This result is of particular interest for all parents who are concerned about any potential negative consequences for their preterm infant when refusing trial participation.

Our study is strengthened by the following factors: (1) using appropriate comparison groups in which the trial patients differed from the non-trial patients in exposure to three doses of NaCl 0.9% only; (2) equally distributing at baseline for the most important prognostic factors; (3) adjusting our primary outcomes for the parental SES, maternal age, and infants' age at follow-up, and (4) evaluating the trial effect on a long-term basis. The main limitations of this study relate to its retrospective design and to the lack of complete datasets at the 2-year follow-up examination for some infants. This loss to follow-up bias is a major drawback of all long-term follow-up studies and results in unclear consequences for the remaining study population [17–19]. In addition, there is increasing evidence that 2-year outcome assessments in very preterm infants might be limited in predicting later cognitive functions [20].

In conclusion, we found additional evidence to support our hypothesis that trial participation of very preterm infants allocated to a placebo group is associated with an equivalent long-term outcome compared to concurrent nonparticipants. In addition to and regardless of any benefits, parents can be told that neither trial participation nor non-participation is related to any harmful or negative effects given that the infants within and outside the trial receive the same or similar treatment.

Acknowledgment

C.M.R. was supported by a Swiss National Science Foundation (SNSF) Career Award for Medical Scientists (33CM30_140334).

Disclosure Statement

The authors have no conflicts of interest to disclose.

References

1 Streptomycin Treatment of Pulmonary Tuberculosis – A Medical Research Council Investigation. Br Med J 1948;ii:769–782.
2 Devereaux PJ, Yusuf S: The evolution of the randomized controlled trial and its role in evidence-based decision-making. J Intern Med 2003;254:105–113.
3 Giaccone A, Kirpalani H: Judgment often impossible without randomized trials. Neonatology 2012;102:137–138.
4 Stiller CA: Centralized treatment, entry to trials and survival. Br J Cancer 1994;70:352–362.
5 Peppercorn JM, Weeks JC, Cook EF, Joffe S: Comparison of outcomes in cancer patients treated within and outside clinical trials: conceptual framework and structured review. Lancet 2004;363:263–270.
6 Braunholtz DA, Edwards S, Lilford RJ: Are randomized clinical trials good for us (in the short term)? Evidence for a 'trial effect.' J Clin Epidemiol 2001;54:217–224.
7 Fauchere JC, Dame C, Vonthein R, Koller B, Arri S, Wolf M, et al: An approach to using recombinant erythropoietin for neuroprotection in very preterm infants. Pediatrics 2008; 122:375–382.
8 Reuner G, Rosenkranz J, Pietz J, Horn R: Bayley II. Bayley Scales of Infant Development II. Deutsche Fassung Manual 2008;2.
9 Largo RH, Graf S, Kundu S, Hunziker U, Molinari L: Predicting developmental outcome at school age from infant tests of normal, at-risk and retarded infants. Dev Med Child Neurol 1990;32:30–45.
10 Jones B, Jarvis P, Lewis JA, Ebbutt AF: Trials to assess equivalence: the importance of rigorous methods. Br Med J 1996;313:36–39.
11 Piaggio G, Elbourne DR, Pocock SJ, Evans SJW, Altman DG; CONSORT Group: Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA 2012;308:2594–2604.
12 Schmidt B, Gillie P, Caco C, Roberts J, Roberts R: Do sick newborn infants benefit from participation in a randomized clinical trial? J Pediatr 1999;134:151–155.
13 De Vries LS, van Haastert IC, Benders MJNL, Groenendaal F: Myth: cerebral palsy cannot be predicted by neonatal brain imaging. Semin Fetal Neonatal Med 2011;16:279–287.
14 Koschmann C, Thomson B, Hawkins DS: No evidence of a trial effect in newly diagnosed pediatric acute lymphoblastic leukemia. Arch Pediatr Adolesc Med 2010;164:214–217.
15 West J, Wright J, Tuffnell D, Jankowicz D, West R: Do clinical trials improve quality of care? A comparison of clinical processes and outcomes in patients in a clinical trial and similar patients outside a trial where both groups are managed according to a strict protocol. Qual Saf Health Care 2005;14:175–178.
16 Vist GE, Hagen KB, Devereaux PJ, Bryant D, Kristoffersen DT, Oxman AD: Outcomes of patients who participate in randomised controlled trials compared to similar patients receiving similar interventions who do not participate. Cochrane Database Syst Rev 2007; MR000009.
17 Hille ETM, Elbertse L, Gravenhorst JB, Brand R, Verloove-Vanhorick SP; Group DP-1CS: Nonresponse bias in a follow-up study of 19-year-old adolescents born as preterm infants. Pediatrics 2005;116:e662–e666.
18 Wolke D, Söhne B, Ohrt B, Riegel K: Follow-up of preterm children: important to document dropouts. Lancet 1995;345:447.
19 Castro L, Yolton K, Haberman B, Roberto N, Hansen NI, Ambalavanan N, et al: Bias in reported neurodevelopmental outcomes among extremely low birth weight survivors. Pediatrics 2004;114:404–410.
20 Hack M: Poor predictive validity of the Bayley Scales of Infant Development for cognitive function of extremely low birth weight children at school age. Pediatrics 2005;116:333–341.